

# Applying statistical methods to text steganography

NECHTA I., FIONOV A.

*Siberian State University of Telecommunication and Information Sciences*

*Novosibirsk, Russia*

e-mail: [www@inbox.ru](mailto:www@inbox.ru)

## Abstract

This paper presents a survey of text steganography methods used for hiding secret information inside some coverttext. Widely known hiding techniques (such as translation based steganography, text generating and syntactic embedding) and detection are considered. It is shown that statistical analysis has an important role in text steganalysis.

**Keywords:** Steganography, steganalysis, linguistic stegosystem, statistical attacks.

## Introduction

Steganography is the art and science of writing hidden messages in such a way that no one, apart from the sender and intended recipient, suspects the existence of the message. In steganography, it is very important to find a good coverttext suitable for embedding hidden messages. This paper provides a basic introduction to steganography and steganalysis, with a particular focus on text steganography. Information hiding techniques are discussed, providing motivation for moving toward text steganography and steganalysis. We will show some of the problems inherent in text steganography as well as issues with existing solutions.

## 1 Steganography

In 1984, Gustavus Simmons illustrated what is now widely known as the prisoners' problem: Let us consider Fig1. two accomplices in a crime, Alice and Bob, are arrested in separate cells. They want to coordinate an escape plan, but their only means of communication is by way of messages conveyed for them by Wendy the warden. Should Alice and Bob try to exchange messages that are not completely open to Wendy, or ones that seem suspicious to her, they will be put into a high security prison. Alice and Bob will have to deceive the warden by finding a way of communicating secretly in the exchanges. It can be done such way: Alice gets any text (coverttext) which does not arise the warden suspicion and embeds (using steganographic method) secret message into it. Then she sends the coverttext with message to Bob. This coverttext is available to both Warden and Bob, but it contains different information to Wendy than to Bob.

Many types of coverttexts are based on data having redundancy, such as video, audio, or image files. In this article we'll discuss one of the areas of steganography, which uses text files as a coverttext. A file with coverttext is called container.

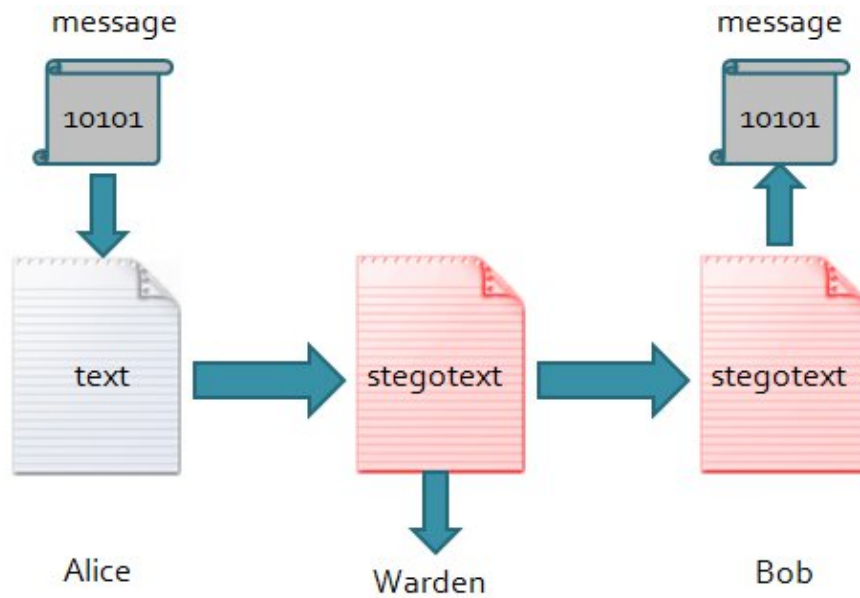


Figure 1: Basic steganography protocol

## 2 Information hiding methods

Existing methods of embedding secret messages in the text data could be divided into three groups:

1. **Syntactic methods** The early methods of text information hiding are based on the physical formatting of text. One of such methods, for example, proposed in [1], uses the extra space between words. One space means that the transmitted information bit is “0”, and two spaces mean “1”. This technique is widely used in HTML files (web pages) because space presence does not affect on the web page appearance. The disadvantage of this method is easy detectability, extra spaces are not used in text. It is possible to use special characters instead of spaces which do not appear in commonly used text editors.

Another method proposed in [1] uses a syntax error when writing words such as:  
*“This is the end”*  
*“This iz the end”*

The second version has a misprint. The presence of errors in certain words (in particular “is”) means that the transmitted information bit is “0”, and errors absence means that bit is “1”. Thus, there is a transfer of information in the text. This method is not easily detectable, because some errors may occur in the message.

2. **Semantic methods** This group includes Tyrannosaurus Lex(T-Lex), published in [2], which uses the replacement words in the sentence on their synonyms, for example:

	<b>excellent</b>		<b>city</b>
	(0) decent		(0)metropolis
Tobolsk is a	little		
	(1) fine		(1) town

Message embeded by synonym selection. Sentence “*Tobolsk is a decent little town*” contains message - “01”. This method requires a large synonyms dictionary. The below examples illustrate two shortcomings of the T-Lex system. First, it sometimes replaces words with synonyms that do not agree with correct English usage, as seen in the phrase “*soon subsequently dispatched*”. Second, T-Lex also substitutes synonyms that do not agree with the genre and the author style of the given text.

*An invitation to dinner was soon afterwards dispatched*

*An invitation to dinner was soon subsequently dispatched*

*...and make it still better, and say nothing of the bad belongs to you alone.*

*...and make it still better, and say nada of the bad belongs to you alone.*

It is clear that the word “*nada*” does not belong to Jane Austen’s style. Furthermore, the string “*say nada*” of is not part of typical English usage.

There is another approach, proposed in [3], of generating sentence level paraphrases for information hiding. Example:

*The caller identified the bomber as Yussef Attala, 20, from the Balata refugee camp near Nablus.*

*The caller named the bomber as 20-year old Yussef Attala from the Balata refugee camp near Nablus.*

This method has a high degree of secrecy.

### 3. Linguistically-driven generation methods

Let’s consider the method proposed in [4], using a context-free grammar to generate a natural like text.

#### Grammar Rules:

$S \rightarrow ABC$

$A \rightarrow She(0) \mid He(1)$

$B \rightarrow likes(0) \mid hates(1)$

$C \rightarrow milk(0) \mid apples(1)$

This approach produces stegotext that looks similar to the real structure of the original text. It is used a set of grammatical rules to generate stegotext and the choice of each word determines how secret message bits are encoded. The quality of the resulting stegotext directly depends on the quality of the grammar. Today’s most popular stegosystems are Nicetext [5], Texto [6] and Markov-Chain-Based [7], because they have high ratio of the input message size to the generated text size. Also, resulting stegotext looking like natural text but it should be noted that, as usually, such text is meaningless.

The next one approach was proposed at [8]. The key idea is to hide information in a noise than occurs invariably in natural languages transformation. When translation a non-trivial text between a pair of natural languages, there are typically many possible translations. Selecting one of these translations can be used to encode information. For example:

*“Джэк нанес краску на стену”* can be translated as:

*“Jack sprayed paint on the wall”, or “Jack sprayed the wall with paint”.*

### 3 Steganalysis methods

There converse problem of steganography is steganalysis. Its goal is to identify suspected container, determine whether or not they have embedded message in it, and, if possible, recover that message. Statistical attacks are commonly used for stegotext detection. For example, widely known support vector machines (SVMs) [9] are a set of related supervised learning methods that analyze data and recognize patterns, used for classification. There are two types of errors uses for evaluating the steganalysis methods reliability:

*False Positive* errors occur when the method mistakenly flags an natural text as stegotext.

*False Negative* errors occur when the method mistakenly flags stegotext as natural text.

The most easily detectable methods are syntactic because they could be detected by simple analyzer. Presence of double spaces in text might cause suspicion. It was noted earlier that methods of natural-like text generation have one disadvantage — resulting text is meaningless. It requires a human intervention to determine the meaningfulness of the text. However, it is not always possible, because of the large volume of messages transmitted in the network. It is necessary to create automated methods for steganalysis. Nowadays, there are a large number of different steganalysis methods. Let us consider in more detail the following method. A method using semantic shortcomings of methods published in [10]. When you replace the words on their synonyms can break semantic rules, for example:

*“What time is it ?”* Word *“time”* could be replaced as *“period”* or *“duration”* that do not agree with correct English usage. False positive ratio is 38.6%. False negative - 15.1%. Low reliability level makes it difficult to the practical application of this method. In addition, it is requires a lot of time working, and a large database of language rules. Method, proposed in [11], uses word frequency and its variance in the analyzed text. Obtained data and The Support Vector Machine (SVM) used for identify stegotext Nicetext, Textto or Markov-Chain-Based presence when container size more than 5Kb. Sum of errors less than 7.05%.

The most effective steganalysis for Nicetext stegosystem proposed in [12]. In virtue of the concepts in area of information theory, the method uses an information entropy-like statistical variable of words in detected text segment together with its variance as two classification features for SVM. The method was centered on detection for small size text segments estimated in the hundreds in words. The experimental accuracy of

the method on classification of generated text and normal text exceeds 99% when text size is larger than 400 bytes. Even for sentences, the experimental accuracy exceeds 85%.

New method of statistical analysis was suggested in [13]. The compression used for stegotext detection. It is known that an embedding message breaks statistical structure of the container, increasing its entropy. Consequently, the full container will compress worse than empty. Let us consider the example:

A and B are empty and full containers, respectively.

Table 1: Container size before and after compression

container	before compr.	after compr.
A	500	320
B	500	300

Add content of suspected container C into A and B. Compare added content sizes before and after compression.

Table 2: Content size of C

container	before compr.	after compr.
C	50	45
C	50	20

It could be asserted that container C is statistically depend with B, which ensures good compression. This principle used in attack on Text0. Accuracy of detection exceeds 99.98% when text size is larger than 400 bytes.

The next one [14], statistical method used for attack on stegotext Nicetext, Text0, and Markov-Chain-Based. The average length of words, frequency of spaces, letter distribution of the words, first letter of word distribution used as features in SVM classifier. Detection accuracy exceeds 84.42% for text segments larger than 500 bytes.

Table 3: The effectiveness of existing steganalysis methods

stegosystem	400 bytes	1 Kb	5 Kb
Nicetext	99.61%	99.61%	99.61%
Text0	99.98%	99.98%	99.98%
MCB	84.42%	87.61%	99.46%

Analysis of translation-based steganography was published in [15]. This method has improved on a previously proposed linguistic steganalysis method based on word distribution which is targeted for the detection of linguistic steganography like nicetext and texto. The newmethod aims to detect the application of TBS and uses none of the related information about TBS, its only used resource is a word frequency dictionary obtained from a large corpus, or a so called natural frequency dictionary, so it is totally blind. It is known, that stegotext consist less high frequency word then in natural text or translated text. This Method needs to know the machine translator set and coverttext language. The experimental results show that the method accuracy is 87.7% when container size - 20 Kb.

The most effective method was proposed in [16]. It is used two features (words frequency and variance of the words distances as sentences structures) for SVM classification.

Let us consider distances of words in “*What is a web browser?*”.

Table 4: Word distances

	what	web	browser
what	0	1	2
web	1	0	1
browser	2	1	0

It should be noted than the sentences structures of stegotext looks more “noisy” than natural text. The total detection accuracy are 97.65%, 98.88% and 99.69% respectively when the text size is 10Kb, 15Kb and 20 Kb.

## Conclusion

This paper presents a background on the major algorithms of text steganography and steganalysis. It is shown that overwhelming majority of effective steganalysis methods are based on statistical analysis. Most existing embedding methods could be detected with high probability.

## References

- [1] James C. Judge (2009). Steganography: Past, Present, Future. [URL] [http://www.sans.org/reading\\_room/whitepapers/steganography/steganography\\_past\\_present\\_future\\_552](http://www.sans.org/reading_room/whitepapers/steganography/steganography_past_present_future_552)
- [2] Winstein K. Tyrannosaurus lex (1999). [URL] <http://alumni.imsa.edu/~keithw/tlex/>

- [3] Barzilay R., Lee L. (2003). Learning to paraphrase: An Universal approach using multiple-sequence alignment. [URL] <http://www.aclweb.org/anthology/N/N03/N03-1003.pdf>
- [4] Chapman M. , Davida G. . (1997). Hiding the hidden: A software system for concealing ciphertext in innocuous text. *Proceedings of the International Conference on Information and Communications Security*. Vol. **LNCS 1334**. Beijing, China.
- [5] [URL] <ftp://ftp.eenet.ee/pub/FreeBSD/distfiles/nicetext-0.9.tar.gz>
- [6] [URL] <http://www.nic.funet.fi/pub/crypt/steganography/texto.tar.gz>
- [7] [URL] <http://www.eblong.com/zarf/markov/chan.c>
- [8] Grothoff C., Grothoff K., Stutsman R., Alkhutova L. (2009). Translation-based steganography. *Journal of Computer Security*. pp. 269-303.
- [9] Chang C., Lin. C. (2001) LIBSVM : a library for support vector machines.(2001).[URL] <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
- [10] Taskiran C., Topkara U., Topkara M., Delp E. (2006). Attacks on Lexical Natural Language Steganography Systems // *Proceedings of the SPIE Security, Steganography, and Watermarking of Multimedia Contents VIII*. Vol. **6072**. pp. 97-105.
- [11] Chen Z., Huang L., Yu Z., Zhao X., Zheng X. (2008). Effective Linguistic Steganography Detection. *IEEE 8th International Conference on Computer and Information Technology Workshops*. pp. 224-229.
- [12] Meng P., Huang L., Chen Z., Yang W., Li D. (2008). Linguistic Steganography Detection Based on Perplexity. [URL] [http://ieeexplore.ieee.org/xpl/freeabs\\_all.jsp?tp=&arnumber=5089098&isnumber=5089035](http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?tp=&arnumber=5089098&isnumber=5089035)
- [13] Нечта И.В. (2010). Эффективный метод стегоанализа базирующийся на сжатии №1. стр. 50-55.
- [14] Hao Y., Xianbin C. (2010). Linguistic Steganalysis Based on Meta Features and Immune Mechanism *Chinese Journal of Electronics*. Vol. **19**, pp. 661-666.
- [15] Meng P. (2010). Analysis and Detection of Translation-Based Steganography. *Lecture Notes in Computer Science*. Vol. **6387/2010**, Springer-Verlag. pp. 208-220.
- [16] Chen Z. (2011). Blind linguistic steganalysis against translation based steganography. *Lecture Notes in Computer Science*. Vol. **6526/2011**, Springer-Verlag, Berlin. pp. 251-265.